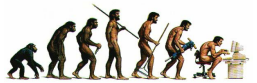
Human-Computer Interaction

Session 9
Natural Dialog Interaction

Evolution of user interfaces



Year	Paradigm	Implementation
1950s	None	Switches, punched cards
1970s	Typewriter	Command-line interface
1980s	Desktop	Graphical UI (GUI), direct manipulation
1980s+	Spoken Natural Language	Speech recognition/synthesis, Natural language processing, dialogue systems
1990s+	Natural interaction	Perceptual, multimodal, interactive, conversational, tangible, adaptive
2000s+	Social interaction	Agent-based, anthropomorphic,social, emotional, affective, collaborative

Overview: machines as…

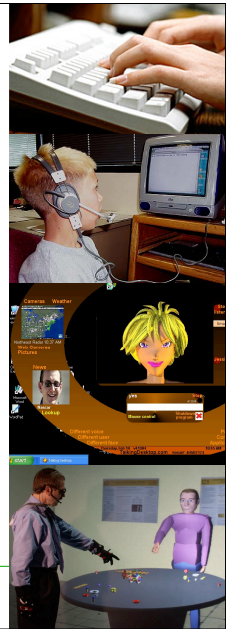
tools ➜ operate

smart tools ➜ instruct

Dialogue Systems

interactive interlocutors ➜ converse

companions ➜ collaborate

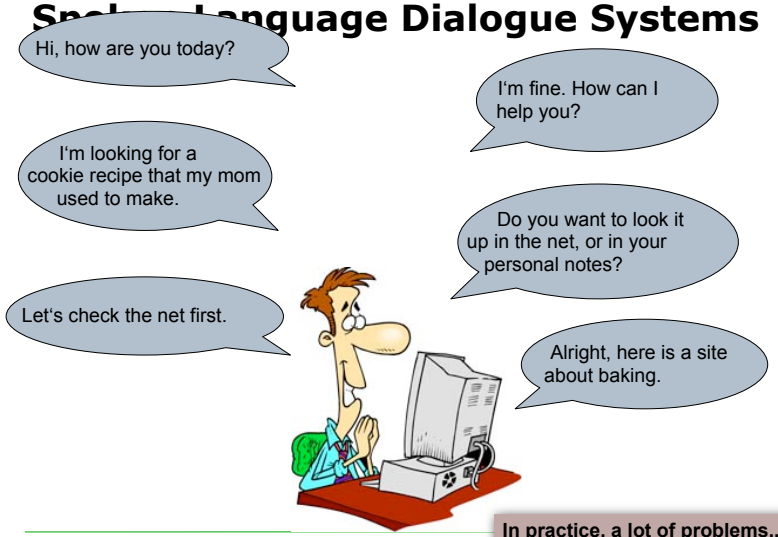


Spoken Language Dialogue Systems (SLDS)

A system that allows a user to speak his queries in natural language and receive useful spoken responses from it

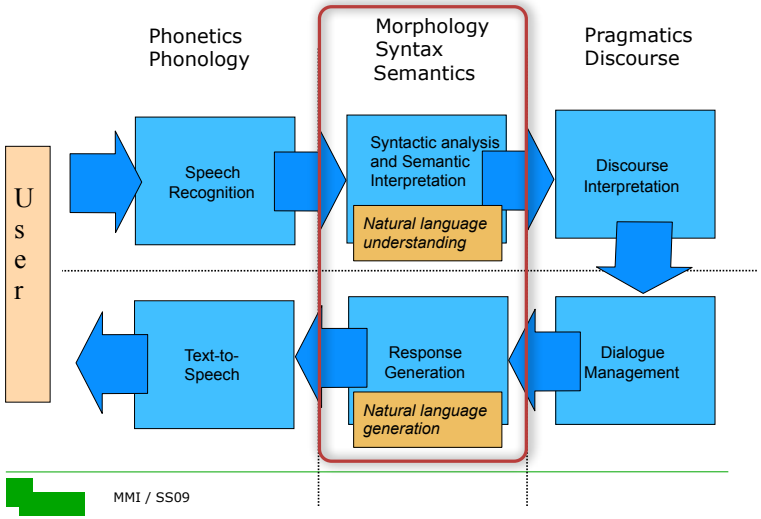
Provides an interface between the user and a computer-based application that permits spoken interaction in a "relatively natural manner"

Spoken Language Dialogue Systems



In practice, a lot of problems...

Classical SLDS structure

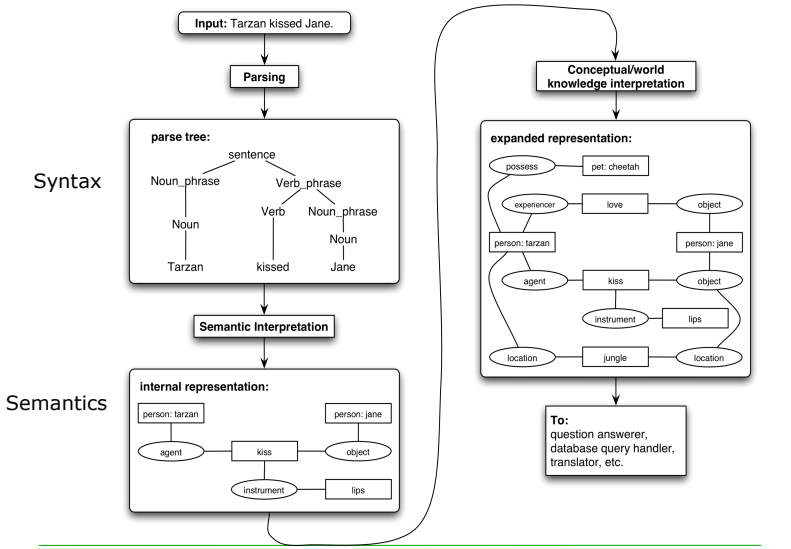


Natural language understanding

Tree classical steps:

1. **Syntax analysis/parsing:**
   - Determine sentence structure from sequence of words
2. **Semantic interpretation/understanding:**
   - Determine word meanings and the overall meaning of their composition in the sentence
3. **Discourse interpretation/pragmatic analysis:**
   - Use context information to complete and disambiguate sentence meaning
   - Determine intention behind the sentence

Allen J. (1995) Natural Language Understanding.

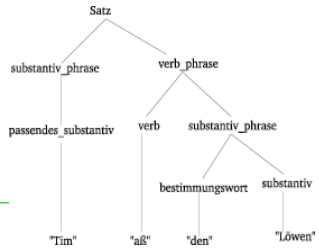


Syntax analysis - parsing

Ziel: Baumartige Zerlegung des sprachlichen Ausdrucks in seine Komponenten gemäß einer Grammatik

```
PARSE ("the dog is dead", G):
        [S: [NP: [Article: the][Noun: dog]]
        [VP: [Verb: is][Adjective: dead]]]
```

- Grammatik: Formale, endliche Beschreibung der *Struktur* aller Elemente einer (oft unendlichen) Sprache
- Parsing = Suchen nach einer möglichen Ableitung eines Satzes in einer Grammatik → Ableitungsbaum
- Beispiel für „Tim aß den Löwen"



Semantic interpretation

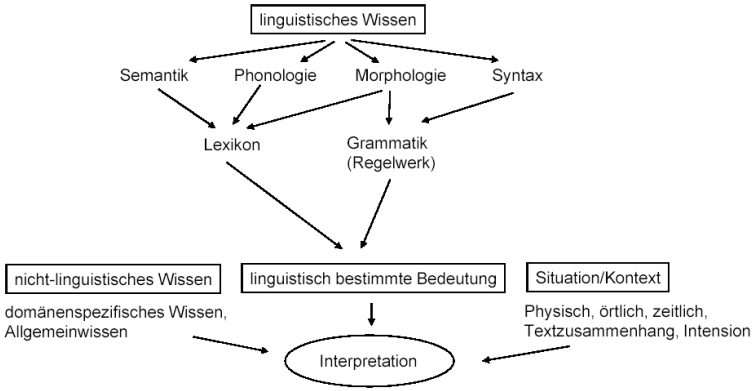
- Aufgabe: *Bedeutungsrekonstruktion*
  - Was ist die *Bedeutung* von *„Er beginnt um zwei im Raum V2-122."* ?
- Unterscheide:
  - **Semantisches Potential**: Linguistisch bestimmte Bedeutung, lässt sich allein mit linguistischem Wissen ermitteln

$$Begin(e,t,l) \wedge Event(e) \wedge Time(t) \wedge Location(l)$$
$$\wedge Equal(t,2) \wedge Room(l,V2-122,?b)$$

  - Aktueller **semantischer Wert**: Volle Interpretation unter Anwendung nicht-linguistischens Wissens (Kontext, Domäne, Welt):

$$Begin(e,t,l) \wedge Event(e) \wedge Time(t) \wedge Location(l)$$
$$\wedge Equal(t,2) \wedge Room(l,V2-122,?b)$$
$$\wedge Talk(e,s,l) \wedge Proffessor(s,Cambridge)$$
$$\wedge Name(s,Steven-Hawking) \wedge Building(b,Uni-Bielefeld) \wedge \ldots$$

Semantic interpretation



Semantic interpretation

Ziel: Bestimmung des semantischen Potenzials

- Umformung des *Parse*-Baumes in eine *interne Repräsentation* (z.B. Prädikatenlogik, Frames, …)
- Zwei wesentliche Schritte:
  1. **Lexikalische Semantik**: Bestimmung der Bedeutung einzelner Wörter
     - Probleme: Homonymie, Polysemie (bank/bank), Synonyme (big/large), Antonyme (boy/girl, hot/cold)
     - Resourcen, z.B. *WordNet* (http://wordnet.princeton.edu/)
  2. **Satzsemantik**: Konstruktion der Gesamtbedeutung aus den Einzelbedeutungen (*kompositionelle* Semantik),
     - häufig anhand des *Parse*-Baums, erweitert mit sem. Kategorien (Name, Aktionsbeschreibung, etc.) *syntaktisch-semantisches Parsing*

Discourse interpretation

Ziel: Von Satzsemantik zu Text-/Diskurssemantik/sem. Wert

- Nötige Wissensquellen (über ling. Wissen hinaus):
  - Domänenwissen (banking transaction)
  - Diskurswissen (satzübergreifend)
  - Weltwissen (*Common-sense knowledge*, Situationswissen)

- Beispiel:

U: I would like to open a fixed deposit account.
S: For what amount?
U: Make it for 8000 Rupees.
S: For what duration?
U: What is the interest rate for 3 months?
S: Six percent.
U: Oh good then make it for that duration.

Discourse interpretation

- Referenzauflösung*: Worauf wird Bezug genommen?*
  - Ellipsen: ausgelassene Wörtern oder Phrasen
  - Anaphern: "John likes that blue car. He buys it."
- Intentionserkennung: Was will der Sprecher?
  - "Do you have the time?" → will die Zeit wissen
  - "When is the last train to London?" → will nach London
- Informationsstruktur: Was ist bekannt und was neu?
- Rhetorische und narrative Struktur: Wie ist der Bezug zum vorher Gesagten?

Benötigt Diskurskontext; siehe unten

Ohne Syntax und Semantik?

Z.B. mit "*keyword-spotting*":

- durchsuchen der Benutzereingabe nach bestimmten Schlüsselworten, z.B. "Wetter", und generieren einer Antwort, die zum Schlüsselwort passt
- Einfach, aber besser skalierbar (grosse Zahl an Regeln)
- Grundlage vieler *Chatterbots*
  - Eliza (Weizenbaum, 1969)
  - ALICE (http://www.alicebot.org/)
  - Jabberwacky.com
  - Anna (www.ikea.de)

- bereits bei einfachen syntaktischen Kniffen überfordert

```
Benutzer: "Ich möchte auf keinen Fall über's Wetter reden!"
Bot: "Gern! Hier in Bielefeld regnet es mal wieder."
```


Natural Language generation (NLG)

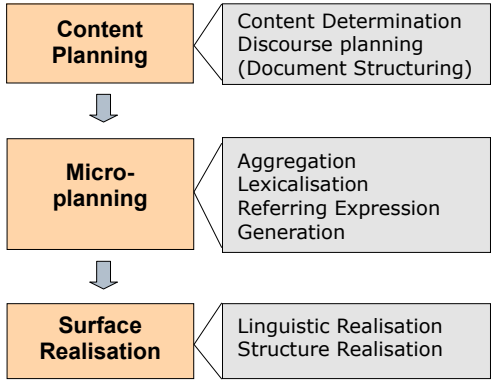
- **Goal**:
  - produce understandable and appropriate output in natural language, along with prosodic information
- **Input**:
  - some underlying non-linguistic representation of information
- **Result**:
  - text to speak, prosodic information
- Knowledge sources required:
  - linguistic knowledge (of language)
  - domain and world knowledge

E. Reiter & R. Dale (2000) *Building Natural Language Generation Systems*. Cambridge University Press.

Natural Language Generation

- Simplest generation method is using *templates*, mapping representation straight to text template (with variables/ slots to fill in).
  - loves(X, Y) → X "loves" Y
  - gives(X, Y, Z) → X "gives the" Y "to" Z
- Templates are very rigid, much more to NLG in general..
  - Consider "John eats the cheese. John eats the apple. John sneezes. John laughs."
  - Better: "John eats the cheese and apple, then sneezes. He then laughs."
- Getting good *style* involves working out how to map many facts to one sentence, when to use pronouns, when to use connectives like "then" etc.

Tasks in NLG



- what to say, how to order and structure it
- how to break it up into sentences and words, how to refer to objects
- How to express things in terms of grammatically correct sentences

1. Content Planning

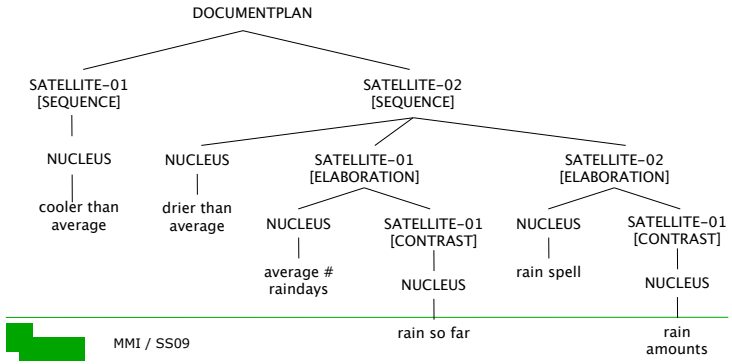
Goals:
- determine *what* information to communicate (content)
- determine *structure* of this information to make a coherent text/discourse

Results: *messages*, predefined data structures that…
- correspond to informational elements (units)
- collect underlying data in ways convenient for ling. expression

- Essentially, a domain-dependent expert-system task
- Common approaches:
  1. based on observations about common utterance structures
  2. based on reasoning about discourse coherence and the purpose of the utterance

Content plan (aka. document plan)

- Tree structure with messages at its leaf nodes
- Rhetorical Structure Theory (RST): distinction between *nucleus,* the central segment, and the *satellite*, the more peripheral one, and relations between them (e.g. elaboration, contrast, ...)
- Example from *WeatherReporter* system (Reiter et al.):



2. Microplanning

Goal:

- convert a content plan into a sequence of sentence or phrase specifications

Tasks:

- **Aggregation** via *conjunction, ellipsis, or embedding*
  - `Heavy rain fell on the 27th and [] on the 28th.`
- **Lexicalisation***:* choosing word lemmas
- **Reference**: how to refer to entities
  - `initially: full name, relate to salient object, specify location`
  - `subsequently: Pronouns, definite NPs, proper names, possibly abbreviated`

3. Surface realisation

Goal:

convert text specifications into actual text

Purpose:

hide peculiarities of English (or whatever the target language is) from the rest of the NLG system

Tasks:

- *Structure realisation*
  - Choose markup to convey document structure
- *Linguistic realisation* using specialized grammars
  - Insert function words
  - Choose correct inflection of content words
  - Order words within a sentence
  - Apply orthographic rules